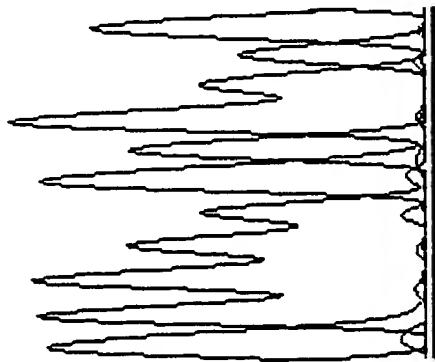


Genetic map of the 11,487 bp ORF on chromosome 1. The map shows the location of various markers and genes. The top part shows the chromosome with markers D13S283, D13S1243, ATP1A1, D13S787, D13S232, D13S1275, and the centromere. The bottom part shows a detailed view of the 11,487 bp ORF with markers M, S, polyA, and AB018273. A scale bar indicates 1 cM and 2 kb.

Figure 1

ACCCTATCA
(SEQ ID NO: 17)



ACCCCATICA
(SEQ ID NO: 18)

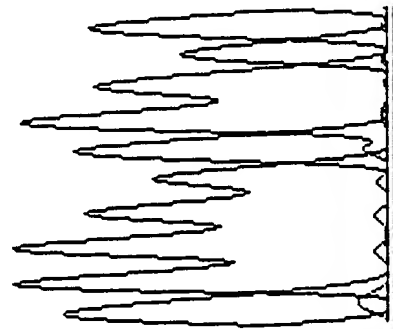
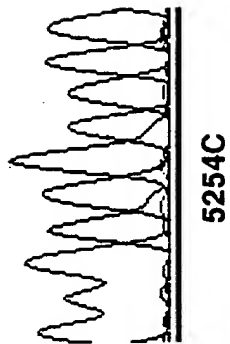


Fig. 2A

AAAGCGACAC
(SEQ ID NO: 19)



AAAGNGACAC
(SEQ ID NO: 20)

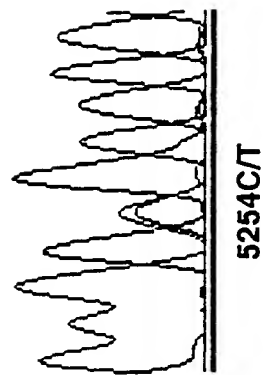


Fig. 2B

Figure 3

Fig. 4A

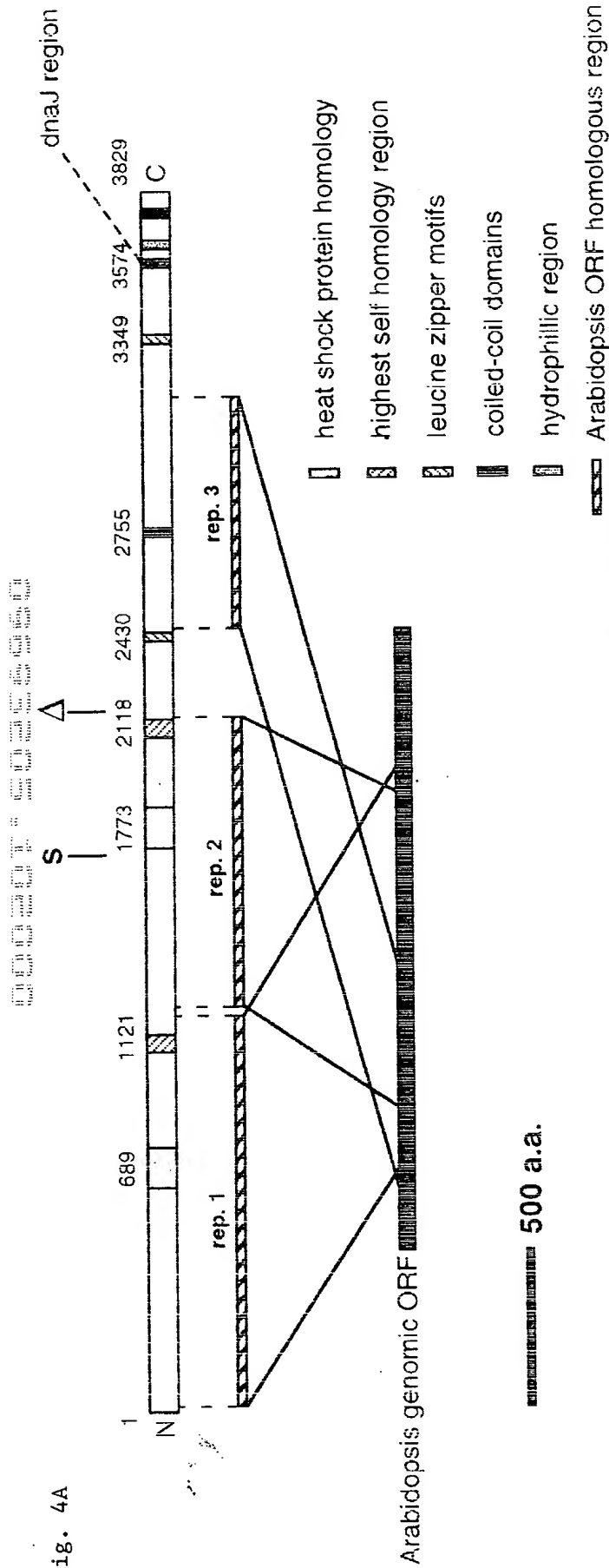
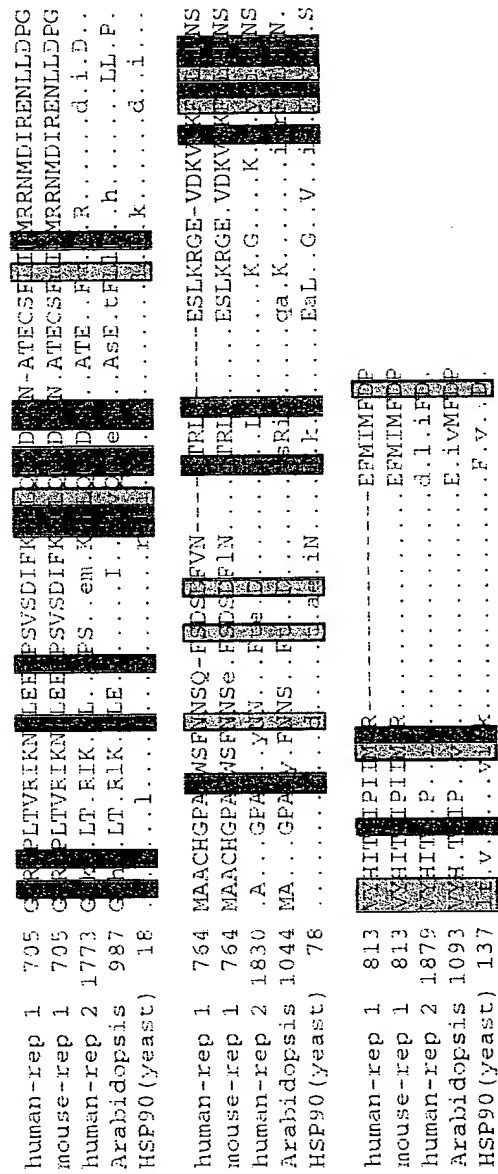


Fig. 4B



(SEQ ID NO: 2)
(SEQ ID NO: 4)

MNTFWPGSEIIVQWYF DENRNHPSVSWL KQVWKNLYIHFSEDLTIFDEMPLIPRTILEE 60
.....S.DKR...L.....L.N.
GQTCVELIRLRIPSLVILDDSEEAQLPEFLADIVQKLGGLVKKLDASIQHPLIKKYIHS 120
D.....V.....T.....I...R...T.....V.....
PLPSAVLQIMEKMPLOKLCNQITSLLPHTKDALRKFLASLTDSSEKEKRIIQELAIFKRI 180
.....I.....I.....A.....T.....T.....
NHSSDQGISSYTKLKGCKVLHHTAKLPADLRSLISVIDSSDEATIRLANMLKIEQLKTTTS 240
.....D.....T.....V.....K.....
CLKLVLDIENAFYSHEEVTQMLWVLENLSSLKNENPNVLEWLTPLKFIQISQEQMVSA 300
...F....G....TQ.....I.....S...D...M....HM...GHV...A.
GELFDPDIEVLKDLFCNEEGTYFPPSVFTSPDILHSLRQIGLKNEASLKEKDVQVAKKI 360
.D.....R...Y...EAC...TI.....S.....R..
EALQVGACPDQDVLLKKAKTLLLVLNKNHTLLQSSEGKMTLKKIKWVPACKERPPNYPGS 420
.....SS.QN....M.....Q.....A.....
LVWKGDLCLNLCAPPDMCDVGHAILIGSSPLVESIHVNLEKALGIFTKPSLSAVLKHFKI 480
.....AA..V.V.....V.....Q..S.....TIN.....T
VVDWYSSKTFSDDEDYQFQHILLEIYGFMHDHLNEGKDSFRALKFPWWVTGKKFCPLAQ 540
.....T.....S.....K.....N.....
VIKPIHDLDLQPYLHNVPKTMARFHLKFKVCGSIEELTSDHISMVIQKIYKSDQDLSEQ 600
...T.....Y.....A.....V.....E...E
ESKQNLHMLNIIIRWLYSNQIPASNPVPIHHSKNPSKLIMKPIHECCYCDIKVDDLND 660
.....M.....Y..R...V.....
LLEDSEVEPIILVHEDI PMKTAEWLKVPCSTRLINPENMGFEQSGQREPLTVRIKNILEE 720
.....
YPSVSDIFKELLQNA DDANATECSFLIDMRNMDIRENLLDPGMAACHGPALWSFNNSQF 780
.....M.....E..
SDSDFVNITRLGESLKRGEVDKVGKFGLGFSVYHITDIPIIMSREFMIMFDPNINHISK 840
.....L.....
HIKDKSNPGIKINWSKQKRLRKFPNQFKPFIDVFGCQLPLTVEAPYSYNGTLFRLSFRT 900
.....R.....A.....
QQEAKVSEVSSTCYNTADIYSLVDEFSLCGHRLLIIFTQSVKSMYLYLKIEETNP SLAQD 960
.....N.....
TVIIKKKSCSSKALNTPVLSVLKEAAKLMKTCSSSNKKLPSDEPKSSCILQITVEEFHHV 1020
.I.....V.P....A.....T.V.....
FRRIADLQSPLFRGPDDPAALFEMAKSGQSKKPSDEL SQKTVECTTWLLCTCMDTGEAL 1080
.....T.....P...D....I.....
KFSLSSESGRRLGLVPCGAVGVQLSEIQDQKWTVKPHIGEVFCYLPLRIKTGLPVHINGCF 1140
....N.....L.H.T.E.....I.....
AVTSNRKEIWKTDTKGRWNTTFMRHVIVKAYLQVLSVLRDLATSGELMDYTYAVWPDPD 1200
.....A.....IG...T.....
LVHDDFSVICQGFYEDIAHGKGKELTKVFS DGSTWVSMKNVRFLDD SILKRRDVGSAAFK 1260
.....K.....R.....M.....Q.K.....
IFLKYLKKTGSKNLC AVELPSSVKLGFE EAGCKQILLENTFSEKQFFSEVFFPNIQIEIA 1320
.....A.....
ELRDPLMIFVLNEKVDEFSGVLRVTPCIPCSLEGHPLVLP SRLIHPEGRVAKLFDIKDGR 1380
.....N.....L.....I.....V.....T.....
FPYGSTQDYLNPIILIKLVQLGKAKDDILWDDMLERAVSVAEINKSDHVAACLRSSILLS 1440
.....M.....E.....A.....

Figure 5A

LIDEKLRKIRDPRAKDFAAKYQTI RFLPFLTKPAGFSLDWKGN SFK PETMFAATDLYTAEH 1500
.....K.....P.....E.....I.....Y
QDIVC.....ILNENSHSFRGCGSVSLAVKEFLGLLKKPTVD.....QLKEVAKSVDDGITL 1560
.....Q.....
YQENITNACYKYLHEALMQNEITKMSIIDKLPFSFILVENAYVDSEKVSFHLNFEAAPY 1620
.....VL.....MA.AT..E.....C.....V..E.....
LYQLPNKYKNNFREL FETVGVRQSCTVEDFALVLESIDQERG TKQITEENFQLCRRRISE 1680
.....S.....F.....K.....
GIWSLIREKKQEFCEKNYKILLPDTNMLLPKSLCYND CPWIKVKDTTVKYCHADIPR 1740
.....R.....L.....S.....
EVAVKLGAVPKRHKALERYASNVCFTTLGTEFGQKEKLT SRIKSILNAYPSEKEMLKELL 1800
.....I.....I.....A.....
ONADDAKATEICEVFDPROHPVDRIFDCKWAPLQGPALCVYNNQPFTEDDVRGIQNLGKG 1860
.....
TKEGNPKYKTGQYGIGFNSVYHITDCPSFISGNDILC IFDPHARYAPGATSISPGRMFRDL 1920
.....C..H.....G.....V.....
DADFRTQFSDVLDLYLGT HFKLDNCTMFRFPLRNAEMAKVSEISSVPASDRMVQNLLDKL 1980
.....N.....Q.....S.....
RSDGAELLMFLNHMEKISICEIDKSTGALNVLYSVKGKITDGDRLK RKQFHASVIDSVTK 2040
.....A..G.....
KRQLKDIPVQQITYTMDTEDSEGNLTTWLICNRSGFSSMEKVSKSVISAHKNQDITL FPR 2100
.....
GGVAACITHNYKKPHRAF CFLPLSLETGLPFHVNGHFALDSARRNLWRDDNGVGRSDWN 2160
.....
NSLMTALIAPAYVELLIQLKKRYFPGSDPTLSVLQNTPIHVVKDTLKKFLSFFPVNRLDL 2220
.....
QPDLYCLVKALYNCIHEDMKRLLPVVRAPNIDGSDLHSAVIITWINMSTSNKTRPFFDNL 2280
.....S.....
LQDELQHLKNADYNITTRKTVAENVYRLKHLLEIGFNLVYNCD ETANLYHCLIDADIPV 2340
.....V.....
SYVTPADIRSF LMTFSSPD TNCHIGKLPCRLQQTNLKLFHSLKLLVDYCFKDAEENEIEV 2400
.....V.....S.F..
EGLPLLITLDSVLQTFDAKRPKFLTTYHELIPSRKDLFMTNTLYLKYSNILLNCKVAKVFD 2460
.....I..G.....SV.....
ISSFADLLSSVLPREYKTKSCTKWKNDFASESWLKNAWHFISESVSVKEDQEETKPTFDI 2520
.....N.A.....TD....P..A..V
VVDTLKDWALLPGTKFTV SANQLVVPEGDVLLPLSLMHIAVFPNAQSDKV FHALMKAGCI 2580
I..I.....TS.....I.....
QLALNKICSKDSAFVPLLSCHTANIESPTSILKALHYMVQTSTFRAEKL VENDORFEALLMY 2640
.....L.....D..A....V.....T..M.....
FNCNLNHLMSQDDIKILKSLPCYKSISGRYVSIGKFGTCYVLT KSIPSAEVEKWTQSSSS 2700
.....S.....M..A.....
AFLEEKIHLKELYEVIGCVFVDDLEVYLKHLLPKIENLSYDAKLEHLIY LKNRL SSAEEL 2760
.....V.....L.....A.I..P
SEIKEOLF EKLESLLIIH DANSRLKQAKHFYDRTVRVF EVMLPEKLFIPNDFFKKLEQLI 2820
.....N.....KE.....V..
KPKNHVT FMTSWVEFLRNIGLKYILSQQQLLQFAKEISVRANTENWSKETLQNTVDILLH 2880
....QAA.....A.....S.....

Figure 5B

HIFQERMDDLSSGNFLKELSLIPFLCPERAPAEFIRFHPQYQEVNGTLPLIKFNGAQVNP 2940
Y.....
 FKQCDVLQLLWTSCPIPEKATPLSIKEQEGSDLGPEQLEQVLNMLNVNLDPPDKVIN 3000
A.....
 NCRNICNITTLDEEMVKTRAKVLRISIYEFLSAEKREFRQLRGVAFVMVEDGWKLLKPEE 3060

 VVINLEYESDFKPYLYKLPLELGTFFHQLFKHLGTEDIISTKQYVEVLSRIFKNSEKQLD 3120
A.....S.....
 PNEMRTVKRVVSGLFRLQNDVSVKVRSDLENVRDLALYLPQDGRILVKSSILVFDDAPHY 3180
K.....A.....K.....
 KSRIQGNIGVQMLVDLSQCYLGKDHGFHTKLIIMLFPQKLRPRLSSILEEQLEETPKVC 3240

 QFGALCSLQGRQLQLLSSEQFITGLIRIMKHENDNAFLANEKAIKALREGLKVSCF 3300

 EKLQTTLRVKGFNPIPHSRSETFAFLKRFGNAVILLYIQHSDSKDINFLLALAMTLKSAT 3360

 DNLI SDTSYLT IAMLGCNDIYRIGEKLDLGVKYDSSEPSKLELPMPGTPIPAEIHYTLLM 3420
S.....
 DPMNVFYPGEYVGYLVDAEGGDIYGSYQPTYTYAIIVQEVEREDADNSSFLGKIYQIDIG 3480
T.....
 YSEYKIVSSLDLYKFSRPEESSQSRDSAPSTPTSPTEFLTPLGLRSIPPLFSGRESHKT-S 3540
D...N...T.....K....SP..
 SKHQSPKKLVNSLPEILKEVTSVVEQAWKI [REDACTED] 3600
 T..H..R....A.....
 [REDACTED] STSASRFQSDKYSFQRFYT SWNOEATSHK 3660

 SERCOONKEKCPPSAGQYSQRFFVPPTFKSVGNPVEARRWLRQARANFSAARNDLHKNA 3720
S.....
 NEWVCFKCYLSTKLALIAADYAVRGKSDKDVKP TALAOKIEEYSOOLEGLTNDVHTLEAY 3780

 GVDSLKTRYPDLLPFPQIPNDRFTSEVAMRVMECTACIIKLENFMQQKV 3830
I.....

Figure 5C

Table 1 ESTs identified by sample-sequencing of the ARSACS critical interval

BAC clone	GenBank #	UniGene	Identity ^a	Tissue Source
235_1_20	AA987300	U15129092	221/230	neuroendocrine lung carcinoids
235_1_20	AA476635*		249/296	total fetus
235_1_20	AI351876*		272/335	melanocyte, fetal heart, pregnant uterus (pool)
235_1_20	W25994	U15163732	447/464	retina
235_1_20	AI377467	U15163732	257/263	total fetus
235_1_20	AA601007		488/491	schwannoma tumor
235_1_20	AA324964		214/234	cerebellum
235_1_20, 206_1_1	AA897178 ^b		238/238	fetal lung, testis, and B-cell
235_1_20, 206_1_1	R17106 ^b	U15188560	747/784	brain, adipose tissue
235_1_20, 206_1_1	AB018273 ^b	U15159492	4318/4318	multiple tissue types including brain, CNS, and whole embryo

a: Number of homologous nucleotides between the BAC subclone sequence and the GenBank sequence

b: AA897178, R17106, and AB018273 are all contained within *spastin*.

* AA476635 and AI351876 were homologous to non-overlapping portions of the same M13 subclone sequence.

Table 2 Primers for PCR amplification of the human *spastin* ORF

Primer set	Forward primer	Reverse primer	Product size (bp)
1 (SEQ ID NO: 21)	CCTTCCAGTACTGTGTTATTTGTGAG	CAAGAACTTCCTCAGGGGCATC	(SEQ ID NO: 22) 603
2 (SEQ ID NO: 23)	GATGCA'TCTATATACAACATCCGCT'	GGTGGGAATAGGTTCCTTC	(SEQ ID NO: 24) 581
3 (SEQ ID NO: 25)	AAAAATGAGAA'TCCAAATGTGCT'	GCACTAAGGCTAGGTTTGTGAAG	(SEQ ID NO: 26) 592
4 (SEQ ID NO: 27)	GCTCCTCAC'TTCCTCTTGTG	CGTGAA'TTGGCTTCATGATAA	(SEQ ID NO: 28) 602
5 (SEQ ID NO: 29)	AGCAATCAGATTCCAGCAAGC	GATGGAA'TGTCAGTGATATGG	(SEQ ID NO: 30) 611
6 (SEQ ID NO: 31)	GGGAGAAAGTTGACAAAGTTGGA	CTTTGGTTCATCACATGGGAAG	(SEQ ID NO: 32) 624
7 (SEQ ID NO: 33)	TCCAAAGCATTGAAACACACCT'	CAGTCCCGTAAAGACATCAG	(SEQ ID NO: 34) 631
8 (SEQ ID NO: 35)	CAATGGGTGCTTTGCTGTAC	CGAAGAACTCCCGAGAACTCA	(SEQ ID NO: 36) 620
9 (SEQ ID NO: 37)	GCTGGCTGCAAAACAGATACTAC	GCAAAACATGGTTCAGGGCTTA	(SEQ ID NO: 38) 604
10 (SEQ ID NO: 39)	CAAAACAATCCGCTTCCTTCCAT	ATTATTCGTCCGGCAAAGCTGA	(SEQ ID NO: 40) 651
11 (SEQ ID NO: 41)	TTCCGCGAACTTTTGAACCC	ACACAAGTGTCTGGCCCTTGC	(SEQ ID NO: 42) 625
12 (SEQ ID NO: 43)	GATGCAAAAGCGACAGAAATC	ATACAGCACATTTAGAGCTCCAGT	(SEQ ID NO: 44) 626
13 (SEQ ID NO: 45)	GCATCAGACAGAAATGGTCCAG	GCAATTCAACATA'TGCAGGAG	(SEQ ID NO: 46) 624
14 (SEQ ID NO: 47)	GTGAATGGCCACTTTGCACT	TGATATCAGCAGGGGTCACAT	(SEQ ID NO: 48) 648
15 (SEQ ID NO: 49)	ACCACACGCAAAACAGTAGCA	GCCATGCA'TTCTT'AAGCCCAAG	(SEQ ID NO: 50) 609
16 (SEQ ID NO: 51)	TGACATTTCCAGCTTTGCTGA	AGCGGCCACTGATGGA'TTTAT	(SEQ ID NO: 52) 631
17 (SEQ ID NO: 53)	AAATGATTTTGAGGCACTTTTG	TTCCACCCAGGATGTCATAAA	(SEQ ID NO: 54) 609
18 (SEQ ID NO: 55)	ACAGTAGACTAAAGCAAGCAAGC	ATCAAGAGGAGGATCCAGGT'T	(SEQ ID NO: 56) 645
19 (SEQ ID NO: 57)	CATCCTGCCCTATTC'TCCAG	TAAAGCGCAAGG'TCTCGTACA	(SEQ ID NO: 58) 618
20 (SEQ ID NO: 59)	TGAGGGCAAAACAATAGATCC	TC'TGCTGTGGGGAATAGGATTT	(SEQ ID NO: 60) 612
21 (SEQ ID NO: 61)	GCAAAAGCCCTAAGAGAAAGGATT	TGCTTTGAGAGCTTTCCTCAG	(SEQ ID NO: 62) 647
22 (SEQ ID NO: 63)	TGAAAGAGAGAAGATGCTGACAATTC	GTAAGTCTGTCCGGCTGAAGG	(SEQ ID NO: 64) 654
23 (SEQ ID NO: 66)	CATCCCGATTTCAGTCAGACA	TTCTGTCTACAACACATTCAGA	(SEQ ID NO: 66) 638

Figure 7

LOCUS AF193557 11493 bp DNA ROD
 DEFINITION Mus musculus saccin gene, complete cds.
 ACCESSION AF193557
 VERSION AF193557.1 GI:6907043
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 11493)
 AUTHORS Engert, J.C., Berube, P., Mercier, J., Dore, C., Lepage, P., Ge, B.,
 Bouchard, J.P., Mathieu, J., Melancon, S.B., Schalling, M.,
 Lander, E.S., Morgan, K., Hudson, T.J. and Richter, A.
 TITLE ARSACS, a spastic ataxia common in northeastern Quebec, is caused
 by mutations in a new gene encoding an 11.5-kb ORF
 JOURNAL Nat. Genet. 24 (2), 120-125 (2000)
 MEDLINE 20120709
 REFERENCE 2 (bases 1 to 11493)
 AUTHORS Engert, J.C., Berube, P., Dore, C., Lepage, P., Ge, B., Hudson, T.J. and
 Richter, A.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) Genome Centre, Montreal General Hospital,
 1650 Cedar Ave., Montreal, QC H3G 1A4, Canada
 FEATURES Location/Qualifiers
 source 1..11493
 /organism="Mus musculus"
 /db_xref="taxon:10090"
mRNA <1..>11493
 /product="saccin"
CDS 1..11493
 /note="molecular chaperone"
 /codon_start=1
 /product="saccin"
 /protein_id="AAF31263.1"
 /db_xref="GI:6907044"
 /translation="MNTFWPGRELVVQWYPFSEDKRHPSLSWLKMVWKNLYIHFSEDL
 TLFDEMPLIPRTLLNEDQTCVELIRLRIPSVVILDDTEAQLPEFLADIVQKLGGIVL
 KRLDTSIQHPLVKKYIHSPLPSAILQIMEKIPLQKLCNQIASLLPTHKDALKRFLASL
 TDTSEKEKRRIIQELTIFKRINHSSDQGISSYTKLKGCKVLDHTAKLPTDLRLSVSVID
 SSDEATIRLANMLKIEKLKTTSCCLKFVLKDIGNAFYTQEEVTQLMLWILENLSSLKNE
 NSNVLDWLMPLKFIHMSQGHVVAAGDLFDPDIEVLRDLFYNEEEACFPPTIFTSPDIL
 HSLRQIGLKNESLKEKD VVQVARKIEALQVSSCQNQDVLMMKAKTLLLVLNKNQTLL
 QSSEGKMAKKIKWVPACKERPPNYPGSLVWKGDLNLCAPPDMCDAAHAVLVGSSLP
 LVESVHVNLQALSIFTKPTINAVLKHFKTVVDWYTSKTFSDDEDYYQFHILLEIYGF
 MHDHLSEGKDSFKALKFPVWWTGKNFCPLAQAVIKPTHDLDLQPYLYNVPKTMKAFHQ
 LFKACGSIEELTSDHISMVIQKVYLKSDQELSEESKQNLHMLNIMRWLYSNQIPAS

Figure 8A

PNTVPVIYHSRNPSKLVMPKPIHECCYCDIKVDDLNDLLEDSEPIILVHEDIPMKTAE
WLKVPCLSTRLINPENMGFEQSGQREPLTVRIKNILEEYPSVSDIFKELLQNADDANA
TECSFMIDMRNMDIRENLLDPGMAACHGPALWSFNNSEFSDSDFLNITRLGESLKR
EVDKVGKFGFGNSVYHITDIPIIMSREFMIMFDPNINHISKHIKDRSNPGIKINWSK
QQKRLRKFPNQFKPFIDVFGCQLPLAVEAPYSYNGTLFRLSFRTQQEAKVSEVSSTCY
NTADIYSLVDEFSLCGHRLLIFTQSVNSMYLKYLKIEETNP SLAQDTIIKKKVCPSK
ALNAPVLSVLKEAAKLMKTCSSSNKKLPTDVPKSSCILQITVEEFHHVFRRIADLQSP
LFRGPDDDPATLFEMAKSGQSKKPSDELQKTVDCCTTWLICTCMDTGEALKFSLNESG
RRLGLVPCGAVGVLLHETQEQQWTVKPHIGEVFCYLPLRIKTGLPIHINGCFAVTSNR
KEIWKTDTKGRWNTTFMRHVIVKAYLQALSVLRLDAIGGELTDYTTYAVWPDPLVHD
DFSVICKGFYEDIAHGKGKELTRVFS DGSMWVSMKNVRFLDDSILQRKDVGSAAFKIF
LKYLKKTGSKNLCAVELPSSVKAGFEEAGCKQILLENTFSEKQFFSEVFFPNIQEIEA
ELRDPLMNFVLNEKLDEFSGILRVTPCVPCSLEGHPLVLP SRLIHPEGRVAKLFDTKD
GRFPYGSTQDYLNPIILIKLVQLGMAKDDILWDDMLERAESVAEINKSDHAAACLRSS
ILLSLIDEKLKIKDPRAKDFAAKYQTIPFLPFLTKPAGFSLEWKGN SFKPKETMFAATD
IYTAEQDIVCLLQPI LNENSHSFRGCGSVSLAVKEFLGLLKKPTVDLVINQLKQVAK
SVDDGITLYQENITNACYKYLHEAVLQNE MAKATII EKLPFCFILVENVYVESEKVS
FHLNFEEAAPYLYQLPNKYKNNFREL FESVGRQSFTVEDFALVLESIDQERGKKQITE
ENFQLCRRIISEGIWSLIREKRQEFCEKNYKILLPD TNLLLLPAKSLCYND CPWIKV
KDSTVKYCHADIPREVAVKLGAIPKRHKALERYASNICFTALGTEFGQKEKLT SRIKS
ILNAYPSEKEMLKELLQNADDAKATEICFVFDPRQHPVDRI FDDKWAPLQGPALCVYN
NQPFTEDDVRGIQNLGKGTKEGNPCKTGHYGIGFNSVYHITDCPSFISGNDILGIFDP
HARYAPGATSVSPGRMFRDLDA DFRTQFSDVLDLYLGNHFKLDNCTMFRFPLRNAEMA
QVSEISSVPSSDRMVQNLLDKLRSDGAELLMFLNHMEKISICEIDKATGGLNVLYSVK
GKITDGDRLKRKQFHASVIDSVTKKRQLKDIPVQQITYTMDTEDSEGNLTTWLICNRS
GFSSMEKVSKSVISAHKNQDITLFP RGGVAACITHNYKKPHRAFCFLPLSLETGLPFH
VNGHFALDSARRNLWRDDNGVGRSDWNNSLMTALIAPAYVELLIQLKKRYFPGSDPT
LSVLQNTPIHVVKDTLKKFLSFFPVNRLDLQPDLYCLVKALYSCIHEDMKRLLPVVRA

Figure 8B

PNIDGSDLHSAV IITWINMSTSNKTRPFFDNLQDELQHLKNADYNITTRKTVAENVY
RLKHLLEIGFNLVYNCDETANLYHCLVDADIPVSYVTPADVRSFLMTFSSPDTNCHI
GKLPCRLLQQTNLKLFHSLKLLVDYCFKDAEESEFEVEGLPLLITLDSVLQIFDGKRPK
FLTTHYELIPSRKDLFMNTLYLKYSSVLLNCKVAKVFDISSFADLLSSVLPREYKTKN
CAKWKDNFASESWLKNAWHFISESVSVTDDQEEPKPAFDVIVDILKDWALLPGTKFTV
STSQLVVPEGDVLIPLSLMHIAVFPNAQSDKVFHALMKAGCIQLALNKICSKDSALVP
LLSHTANIDSPASILKAVHYMVQTSTFRTEKLMENDFEALLMYFNCNLSHLMSQDDI
KILKSLPCYKISGRYMSIAKFGTCYVLTKSIPSAEVEKWTQSSSSAFLEEKVHLKEL
YEVLCVVPDDLEVYLKHLPLKIENLSYDAKLEHLIYLNRLASIEEPSEIKEQLFEK
LESLLIHDANNRLKQAKHFYDRTVRVFEVMLPEKLFIPKEFFKKLEQVIKPKNQAAF
MTSWVEFLRNIGLKYALSQQQLLQFAKEISVRANTENWSKETLQSTVDILLHHIFQER
MDLLSGNFLKELSLIPFLCERAPA EYIRFHPQYQEVNGTLPLIKFNGAQVNP KFKQC
DVLQLLWTSCPI LPEKATPLSIKEQEGSDLAPQEQLQVLNMLNVNLDPP LDKVINNC
RNICNITTLDEEMVKTRAKVLRSIYEFLSAEKREFRFQLRGVAFVMVEDGWKLLKPEE
VVINLEYEADFKPYLYKLPLELGT FHQLFKHLGTEDIISTKQYVEVLSRIFKSSEGKQ
LDPNEMRTVKRVVSGFLFKSLQND SVKVRSDLENARDLALYLP SQDGKLVKSSILVFDD
APHYKSRIQGNIGVQMLVDLSQC YLGKDHGFHTKLIMLFPQKLRPRLLSSILEEQLDE
ETPKVCQFGALCSLQGR LQLLSSEQFITGLIRIMKHENDNAFLANE EKAIRLCKALR
EGLKVSCFEKLQTTLRVKGFNP IPHSRSETFAFLKRFGNVILLYIQHSDSKDINFL
ALAMTLKSATDNLISDTSYLIAM LGCNDIYRISEKLD SLGVKYDSSEPSKLELPMPGT
PIPAEIHYTLLMDPMNVFYPGEYVGYLVDAEGGDIYGSYQPTYTYAII VQEVEREDAD
NTSFLGKIYQIDIGYSEYKIVSS LDLYKFSRPDESSQNRDSAPTTP TSPTEFLTPGLR
SIPPLFSGKESHKSPSTKHHS PRKLKVNALPEILKEVTSVVEQAWKLPESERKKIIRR
LYLKWHPDKNPENHDIANEVFKHLQNEINRLEKQAF LDQNADRASRRTFSTSASRFQS
DKYSFQRFYTSWNQEATSHKSERQQQSKEKCPPSAGQTY SQRFFVPPTFKSVGNPVEA
RRWLRQARANFSAARNDLHK NANEWVCFKCYLSTKLALIAADYAVRGKSDKDVKPTAL
AQKIEEYSQQLEGLTNDVHTLEAYGVDSLKTRYPDLLPFPQIPNDRFTSEVAMRVMEC
TACIIIKLENFIQQKV"

Figure 8C

BASE COUNT 3599 a 2281 c 2387 g 3226 t
 ORIGIN

```

1 atgaatacat tctggcctgg tcgagagttg gtgggttcagt ggtatccatt tagtgaagac
61 aaacgtcacc catccctttc atggccttaag atgggtttgga agaatctcta tatacatttc
121 tcggaagatt tgactttatt tgatgagatg ccacttatcc ctagaactct actgaatgag
181 gaccagacgt gtgtggaact catcagactc aggatcccat cagtagtcat tttagatgat
241 gaaactgaag ctcagcttcc agaattctta gcagatattg tacaaaaact tggagggatt
301 gtcctgaaaa gactagatac ctctattcag catccacttg ttaaaaaata cattcattcc
361 ccactcccga gtgctatttt gcagataatg gagaagatac ctctacagaa gttgtgtaat
421 caaatagcat cattacttcc aaccacaaa gatgctctaa ggaagttttt ggccagctta
481 actgatacca gtgaaaaaga gaaaagaata attcaagaat tgacaatatt caaaagaatt
541 aatcactcat cagatcaagg gatttctctc tacacaaaat taaaaggatg taaagttttg
601 gatcataccg ccaagcttcc aacagatcta cggctatcag tttcagtaat agatagtagt
661 gatgaagcca ccattcgttt ggcaaacatg ttgaaaattg aaaaattgaa gactacaagc
721 tgtttaaaagt ttgttttaaa agatattgga aatgcatttt atacacagga agaggtaaca
781 caacttatgc tttggatcct tgagaatcta tcctctctta aaaatgagaa ttcaaatgtg
841 cttgattggg taatgccact aaaattcatt catatgtccc agggacatgt ggtagcagct
901 ggtgatctct ttgatcctga tatagaagta ctaagggatc tcttttataa tgaagaagaa
961 gcttgtttcc cacctacaat ttttacctca ccagatatcc ttcactcttt gagacagatt
1021 ggcttaaaaa atgaatccag tctaaaagaa aaagatgttg tacaagtggc aagaaaaatt
1081 gaagctttac aggtcagttc ctgtcagaat caggatgttc tcatgaagaa agccaaaaca
1141 ctcttactgg tcttgaataa aaaccagaca ctcttgcaat cttctgaagg gaagatggca
1201 ttgaagaaaa tcaaatgggt tccagcctgc aaggaaagac ctccaaatta tcccggttcc
1261 ttagtctgga aaggggatct ctgtaatctt tgtgcacctc cagatatgtg tgatgcggca
1321 catgcagttc tagtaggctc ctcaacttct cttgttgaaa gtgtccatgt gaacctggag
1381 caggcgctca gcatcttcac aaagcctact atcaatgctg tcttaaaaca ctttaaaact
1441 gttgttgact ggtatacttc aaaaaccttt agtgatgaag attactatca gttccaacat
1501 attttgcttg aaatttatgg gttcatgcat gatcatctga gtgaagggaa ggattctttt
1561 aaagccttga agtttccatg gggttggaat ggcaaaaact tttgtcctct tgcccaggtc
1621 gtgataaagc caaccatga tctggatctt cagccttatt tatataatgt gcctaaaacc
1681 atggcaaaat tccaccagct gttcaaggct tgtggctcaa tagaagagt gacatcagat
1741 catatttcca tggtcattca gaaagttaat ctcaaaagtg accaggagt gagtgaagaa
1801 gaaagtaaac aaaatcttca tctcatgttg aatattatga gatggctcta tagcaatcag
1861 attccagcaa gccctaatac accagttcct atttatcaca gcagaaatcc ttccaaactt
1921 gtcatagaag caattcatga atgctgttat tgtgacatca aagttgatga cctcaatgac
1981 ttgcttgaag attcagtggg accaattatc ttggtacatg aagatatacc catgaaaact
2041 gcagaatggc taaaagttcc gtgccttagt acaagactga tcaatcctga aaacatgggg
2101 tttgagcagt cagggcaaag agagcctctt actgtaagga ttaaaaaat tttggaagaa
2161 tacccttccg tgtcagatat ttttaaagag ctacttcaaa atgctgatga tgcaaatgcc
2221 acagaatgca gcttcatgat tgatatgaga aggaatatgg acatacggga aaatctcctg
2281 gaccaggga tggcagcttg tcatggacct gctctgtggg cattcaacaa ttctgaattc
2341 tcagattcag atttcttaaa cataacgagg ttaggagagt ctttaaaaag gggagaagtt
2401 gacaagggtg ggaaatttgg tcttggtttt aattctgtgt accacatcac tgacattccc
2461 atcattatga gcagagaatt tatgataatg tttgatccaa acataaatca tatcagcaaa
2521 cacattaaag atagatcgaa tcctggaatc aaaattaatt ggagtaagca gcagaaaaga
2581 cttaggaagt tccccaacca gttcaaacca tttatagatg tatttggctg tcagttacct
2641 ttggctgttg aagctcctta cagctacaat ggaactcttt tccgactgtc ctttagaaca
2701 cagcaggaag caaaagtga tgaagttagc agtacttgct acaatactgc ggaattttac
2761 tccatagtgg atgaatttag tcttgttggg cacagactta tcatttttac tcagagtgtg
2821 aactcgtatg atttgaaata cttgaaaatt gaagaaacca atcctagctt agcacaagat
2881 acaatcataa ttaagaaaaa agtttgcccc tccaaagcat tgaatgcacc agttttaagt
2941 gttttaaaag aagctgctaa actcatgaag acttgtagca gcagcaacaa gaagcttccc
3001 acggatgtgc caaagtcac ttgcattctt cagatcacag tcgaagaatt ccaccatgtg
3061 tttaggagga ttgctgactt acagtcacca ctatttcgag gtccagatga tgaccagct
3121 actctctttg aaatggctaa atctggccaa tcaaaaaagc catcagatga gttgccacaa
3181 aagacagtag attgtaccac atggcctata tgcacatgca tggatacagg agaagctctc
3241 aagttttcct tgaatgaaag tggagaagaa ttagggctgg ttccttgtgg ggcagtaggg

```

Figure 8D

3301	gttctcttgc	atgaaaccca	ggaacagaag	tggaccgtga	aaccacacat	aggagaagtg
3361	ttttgctatt	tacctctacg	aatcaaaaca	gggttgccaa	ttcacatcaa	tggtgctttt
3421	gctgttactt	caaataaggaa	agaaatctgg	aagacagata	caaaagggtcg	atggaatacc
3481	acattcatga	ggcatgtcat	tgtgaaagct	tacttacaag	ccctcagtgt	cttacgggac
3541	ctagccattg	gtggtgagct	gactgattat	acttactatg	cagtgtggcc	tgatcctgat
3601	ctagttcatg	atgacttctc	tgtgatctgt	aaaggatttt	atgaagacat	tgctcatggg
3661	aaggggaagg	agttgaccag	agtcttctct	gatgggtcta	tgtgggtttc	catgaagaat
3721	gtgaggtttc	tggtgactc	tatacttcaa	aggaaagatg	ttggttcagc	agccttcaag
3781	atatttctga	agtacctcaa	gaaaacagga	tccaaaaacc	tctgtgctgt	tgagcttcct
3841	tcttcagtaa	aagcaggatt	tgaagaggct	ggctgtaaac	agatactgct	ggaaaataca
3901	ttttcagaga	aacagttctt	ttcagaagtc	ttctttccta	atatccagga	aattgaagca
3961	gaacttagag	atcctctgat	gaattttgtc	ctaaatgaaa	aacttgatga	gttctcagga
4021	attcttcctg	ttaccctctg	tgttccttgc	tccttggagg	gccatccttt	ggttttgctt
4081	tcaagattga	tccatcctga	aggacgagtt	gcaaagtatt	ttgatactaa	agatggaaga
4141	ttcccttatg	gttccacaca	ggattacctc	aatcctatta	tcttgattaa	gctcgttcag
4201	ttaggcattg	caaaagatga	tattttgtgg	gatgacatgc	tagagcgtgc	agagtctgta
4261	gctgagatta	ataaaagtga	ccatgctgct	gcctgcttaa	ggagtagtat	tctgctaagc
4321	cttattgatg	agaagctaaa	aataaaggat	cctagagcaa	aggattttgc	tgcaaaatat
4381	caaacaattc	ccttctctcc	atttctaaca	aagccagcag	gtttttcttt	agaatggaaa
4441	gggaacagct	ttaagcctga	aaccatgttt	gcagcaactg	acatttacac	agctgaatat
4501	caagatatag	tctgtctttt	gcaaccaatt	cttaatgaaa	attcccattc	ctttagaggc
4561	tgtggttcag	tgtctttggc	tgtaaggag	tttttgggtt	tactaaagaa	gccaacagtt
4621	gatctggtaa	taaaccagtt	gaagcaagtt	gcaaaatcag	ttgatgatgg	cattacattg
4681	taccaggaaa	atatcaccaa	cgcttgctac	aaatacctcc	atgaagcagt	attgcagaat
4741	gaaatggcca	aggcaacaat	tattgagaag	ctaaagccat	tttgtttcat	tctagttgag
4801	aatgtatatg	ttgagtcaga	aaaggtttct	tttcacttga	actttgaagc	agcaccatac
4861	ctttatcagt	tacctaacaa	gtataaaaa	aatttccgtg	agctttttga	aagtgtgggt
4921	gtgagacagt	catttactgt	tgaagacttt	gccctagtgt	tggagtctat	tgatcaagag
4981	agaggaaaaa	aacaaataac	agaagagaat	tttcagcttt	gccgacgaat	aatcagtgaa
5041	ggcatctgga	gtctcattag	agaaaagaga	caagaatttt	gtgagaaaaa	ttatggcaaa
5101	atattactgc	cagacactaa	cctgctgctg	ctccctgcta	agtcattatg	ctacaatgac
5161	tgtccctgga	taaaagtaaa	ggactccact	gtcaagtatt	gccatgccga	cataccccgg
5221	gaagtagctg	taaaacttgg	tgcaatacca	aagagacata	aagcattaga	aagatatgca
5281	tccaacatct	gtttcacagc	tctaggtaca	gaatttgggc	agaaagaaaa	actgaccagc
5341	agaattaaga	gcattctcaa	tgcctatcct	tcagaaaagg	aaatgctgaa	agagcttctt
5401	caaaatgctg	atgatgcaaa	ggccacagag	atctgctttg	tgtttgatcc	tagacagcat
5461	cctgttgacc	gaatatttga	tgataagtgg	gccccactgc	aagggccagc	actgtgtgtt
5521	tacaacaacc	agccatttac	agaagatgat	gttagaggaa	ttcagaatct	tggaagaggc
5581	accaaagaag	ggaatccttg	caaaacagga	cattatggaa	tcggattcaa	ttccgtttat
5641	catattacag	actgcccttc	ttttatttct	ggcaatgaca	tcctgggtat	ttttgatccc
5701	catgccagat	atgcaccagg	agccacatca	gttagccctg	gacgcatgtt	tagagatttg
5761	gatgcagact	ttagaaccca	gttctcagat	gttctagatc	tgtacttggg	aaaccacttt
5821	aaactggaca	attgtacaat	gtttagattt	cctctgcgta	atgcagagat	ggcacaagtt
5881	tcagaaattt	cttccgttcc	atcatcagac	agaatggtcc	agaatctttt	ggacaagtta
5941	cgtctgatg	gggcagaact	tctaattgtt	ctcaaccaca	tgagaaaaag	atctatttgt
6001	gaaatagata	aggccacagg	aggtctgaat	gtgctctatt	cagtaaaaag	caagatcact
6061	gatggagacc	gattgaaaag	gaagcaattc	cacgcctctg	taattgacag	tgttactaaa
6121	aagagacagc	tcaaggacat	accagttcaa	caaataacct	acactatgga	tactgaggat
6181	tctgaaggaa	atctgaccac	atggctcatc	tgtaatagat	caggattttc	aagtatggaa
6241	aaagtatcca	agagtgtaat	atcagctcac	aagaaccaag	atatcaccct	tttcccacgt
6301	ggtggagtag	cagcctgcat	tactcacaat	tataaaaagc	cccacagagc	cttctgcttt
6361	ctgcctctct	ctttggagac	agggctgcca	tttcatgtga	atggccactt	tgctctagat
6421	tcagccagaa	gaaacttgtg	gcgtgatgat	aatggggttg	gtgttcgaag	tgactggaat
6481	aatagtttaa	tgacagcatt	aatagcacct	gcatatgttg	agttactaat	ccagttaaaa
6541	aaacggtatt	tccctgggtc	tgacccaaca	ttatcagttt	tacagaacac	acccattcat
6601	gtcgtaaaag	acacattaaa	gaagtttctg	tccttctttc	cagttaacag	gctggatctg
6661	cagccggact	tatattgctt	agtaaaagca	ctttacagtt	gcattcatga	agacatgaag

Figure 8E

6721 cgtcttttgc ctgttggttcg ggctccaaat attgatggct cagatttgca ctctgcagtc
6781 ataattactt ggatcaatat gtctacttca aataaaacta gaccattttt tgataactta
6841 ctacaggatg aattacagca ccttaaaaaat gcagattata acatcacaaac tcgaaaaaca
6901 gtcgcagaga atgtctacag actgaagcac ctgctcttag aaattgggtt caacttggtt
6961 tataactgtg atgaaactgc taacctttac cattgccttg tagatgcaga tatccctgtc
7021 agctatgtga ctcttgctga tgttaggtcc ttcttaatga ctttctcttc tctgacact
7081 aattgccata ttgggaagct gccttgctgt cttcagcaga ctaacctaaa actttttcac
7141 agtttaaaac ttttagttga ttactgtttt aaagatgctg aagaaagtga gtttgaagtt
7201 gagggactgc ccctactcat tacactggac agtgtcttgc agatttttga tggtaaacga
7261 cccaagtttc taacaacata ccatgaatta attccatcgc gtaaagactt gtttatgaac
7321 accttatact tgaaatacag tagtgttttg ttgaactgca aagttgcaaa agtgtttgac
7381 atttccagct ttgctgactt actctcttct gtgttgccctc gtgagtacaa gacaaaaaac
7441 tgtgcaaagt ggaaagacaa ttttgccagt gaatcttggc ttaagaacgc atggcatttt
7501 atcagtgaat cagtaagtgt aacggatgat caggaagaac caaagccagc atttgatgtc
7561 attgttgaca tccttaaaga ctgggcattg cttccaggaa caaagttcac tgtgtcaacc
7621 agtcagcttg tggttcctga gggagacgtg ttgattcccc tgagcctcat gcacattgct
7681 gtgttcccaa atgctcagag tgataaggtt tttcacgctc tgatgaaagc tggctgtatt
7741 cagctggctt tgaacaaaat ctgctctaaa gacagcgcat tagttcctct gttgtcatgc
7801 cacacagcaa acatagatag ccctgcaagc atcttgaagg ctgtgcatta tatggttcag
7861 acgtcaacat ttagaactga aaaactaatg gaaaatgact ttgaagcact tttgatgtat
7921 ttcaactgta atttgagtca cttgatgtcc caagatgaca taaaaatttt aaagtcacctc
7981 ccattgctaca aatccatcag tggccgctat atgagcattg caaaatttgg aacgtgctat
8041 gtgcttacca aaagtattcc ttcagctgaa gtggaaaaat ggacacagtc atcctcttcc
8101 gcgtttcttg aagaaaagg tgcatttaaaa gaactctatg aggtgcttgg ctgtgtgcca
8161 gtagatgac tggaggtgta tttgaaacat cttctgccaa aaattgaaaa tctctcttat
8221 gatgcaaagt tggagcacct gatttatctg aagaatagac tggcaagcat cgaggaaccg
8281 tcagagatta aggagcaact ttttgaaaaa ctggaaagct tattgattat ccacgatgcc
8341 aacaactcag taaagcaagc aaaacatttc tatgacagaa ctgtgagagt ttttgaagtt
8401 atgcttctcg aaaaattggt tattcctaag gatttcttta aaaaatttga acaagtaatc
8461 aaacctaaaa atcaagctgc atttatgacg tcctgggtgg aattcttgag aaatattgga
8521 ctgaagtacg cgctctccca gcagcagttg ttacagtttg ccaaggaaat cagtgtgagg
8581 gcaaatacag aaaactggtc taaagaaacc ctgcaaagta cagttgacat ccttctccat
8641 cacatattcc aagaacgaat ggatttggtt tctggaaatt ttctgaaaga actgtcctta
8701 ataccattct tgtgtcctga acgggcccc gctgagtaca ttcggtttca ccctcagtac
8761 caggaggtaa acggaacact tcctcttata aagttcaatg gagcacaagt gaatccaaag
8821 ttcaagcaat gtgatgtact ccagctgctg tggacatctt gccctattct tccagagaaa
8881 gccacaccgt tgagcattaa agaacaagaa ggcagtgacc tcgctccaca ggaacagctt
8941 gaacaagttt taaatatgct taatgttaac ctggaccccc ctcttgataa ggtcattaat
9001 aattgcagaa acatatgcaa cataacaact ttggatgagg aaatggtaaa aactagagca
9061 aaggtcctaa ggagcatata tgaatttctg agtgcagaaa aacgagagtt ccgttttcag
9121 cttcgggggtg tggcctttgt aatggtagaa gacggatgga aacttctgaa gcctgaggaa
9181 gtagtgataa acctggagta tgaggctgat tttaaacctt atctgtacaa gctgccttta
9241 gagcttggca cttttcatca gctgttcaaa catttaggta ctgaagatat catttccact
9301 aagcaatatg ttgaagtgtt aagccgaata ttcaaaagct ctgaaggaaa gcagctagac
9361 cctaatagaa tgcgtacagt taagagagtg gtttctggcc tattcaagag tctacaaaat
9421 gattcagtc aaggtaggag tgacctggag aatgcccggt acctcgact ctacctcca
9481 agccaggatg ggaagtgtgt gaagtcaagc atcttggtgt tcgatgatgc gccacattat
9541 aaaagtagga tccaggggaa tattggcgtg cagatgctag ttgatcttag ccagtgtac
9601 ttaggggaaag accatggatt tcacactaag ctgataatgc tctttctca aaagcttcga
9661 cctcgtctgc tgagcagtat acttgaagag cagcttgatg aggagacccc taaagtgtgc
9721 cagtttggcg cattgtgctc tcttcaggga agactgcagc ttctctgtc ttcagagcag
9781 ttcattcacag gactcattcg aatcatgaag catgaaaatg ataatgctt cctggccaat
9841 gaagaaaaag ccataagact ttgcaaagct ctaagagaag ggctgaaagt ttcctgtttt
9901 gagaagcttc agacaacatt aagggttaaa ggttttaatc ctattcccca tagcaggagt
9961 gaaactttcg cttttctaaa gcgatttggc aatgcagtca tcttgctcta catccaacat
10021 tcagacagca aagacattaa ctttctgcta gccttagcga tgacacttaa atcagcaact
10081 gacaatttga tttctgacac gtcatactta attgctatgc tgggatgcaa tgacatttac

Figure 8F

10141 aggatcagtg agaagcttga cagtttaggg gtgaaatagc actcctctga gccatcaaaa
 10201 ctggaactcc ccatgcctgg cacaccaata cccgctgaga tccattacac actacttatg
 10261 gatccaatga atgtttttta tcctggggaa tatgttggtt accttggtga tgctgaaggt
 10321 ggtgatattc atgggtcata ccagccaaca tacacatacg caattattgt gcaagaagtt
 10381 gaaagagaag atgctgacaa tactagtttc ttaggaaaga tctatcagat cgatattggc
 10441 tacagtgaat ataagatagt cagctctctt gatctgtaca agttctcaag gcctgatgaa
 10501 agtcccaaaa acagagacag tgctcccacc acaccaacaa gccccaccga attcctgact
 10561 cctggtctga gaagcatccc tcctcttttc tctggcaagg agagccacaa gtctccctcc
 10621 accaaacacc attccccag aaagctcaag gtgaatgctt taccagaaat cttaaaagaa
 10681 gtgacatcag tgggtggagca agcttggaag cttccagaat cagagcggaa aaagatcatt
 10741 agacgcttgt atttgaagtg gcaccctgac aaaaatccag aaaatcatga tattgctaatt
 10801 gaagtgttca agcacctgca gaatgaaatc aacagattag aaaaacaggc ttttctggat
 10861 caaatgcag acagagcttc aagaagaaca ttttcaacct ctgcatctcg atttcagtca
 10921 gacaagtact catttcaaag attttacact tcgtggaatc aagaagccac aagtcataaa
 10981 tctgaaaggc aacagcaaag caaagagaaa tgccctcctt ctgctggaca gacatactct
 11041 caaaggttct ttgttcctcc caccttcaag tcagtgggca atccagtggga agcccgagga
 11101 tggttaagac aagccagagc aaacttctca gctgccagga atgaccttca caaaaatgcc
 11161 aatgaatggg tgtgcttcaa gtgttacctt tccaccaagc tggctttgat tgcagccgac
 11221 tatgctgtca gggggaaatc tgataaagat gtaaagccaa ctgcacttgc aaaaagata
 11281 gaggagtaca gtcagcagct ggaaggactg acaaacgatg tgcacacatt ggaagcttat
 11341 ggtgtagaca gcttgaaaac aaggtagcct gatttgcttc cttttccgca gattcccaat
 11401 gacaggttca catctgaggt tgccatgagg gtgatggaat gcactgcctg tatcatcata
 11461 aaacttgaat attttatata acagaagggtg tga

//

Figure 8G

LOCUS AF193556 12793 bp DNA PRI

DEFINITION Homo sapiens saccin (SACS) gene, complete cds.

ACCESSION AF193556 VERSION AF193556.1 GI:6907041KEYWORDS .

SOURCE human.

ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 12793) AUTHORS Engert,J.C., Berube,P., Mercier,J., Dore,C., Lepage,P., Ge,B., Bouchard,J.P., Mathieu,J., Melancon,S.B., Schalling,M., Lander,E.S., Morgan,K., Hudson,T.J. and Richter,A. TITLE ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF JOURNAL Nat. Genet. 24 (2), 120-125 (2000) MEDLINE 20120709

REFERENCE 2 (bases 1 to 12793) AUTHORS Engert,J.C., Berube,P., Dore,C., Lepage,P., Ge,B., Hudson,T.J. and Richter,A. TITLE Direct Submission JOURNAL Submitted (08-OCT-1999) Genome Centre, Montreal General Hospital, 1650 Cedar Ave., Montreal, QC H3G 1A4, CanadaFEATURES Location/Qualifiers source 1..12793 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="13" /map="between D13S232 and D13S292" mRNA 1..12793 /gene="SACS" /product="saccin" gene 1..12793 /gene="SACS" CDS 77..11566 /gene="SACS" /note="molecular chaperone" /codon_start=1 /product="saccin" /protein_id="AAF31262.1" /db_xref="GI:6907042" /translation="

MNTFWPGRELIVQWYPFDENRNHPSVSWLKMVWKNLYIHFSEDL
TLFDEMLIPRTILEEGQTCVELIRLRIPSLVILDDSEEAQLPEFLADIVQKLGGFVL
KKLDASIQHPLIKKYIHSPLPSAVLQIMEKMPLQKLCNQITSLPHTKDALRKFLASL
TDSSEKEKRIIQELAIFKRINHSSDQGISSYTKLKGCKVLHHTAKLPADLRLSISVID
SSDEATIRLANMLKIEQLKTTSCCLKVLKDIENAFYSHEEVTQLMLWVLENLSSLKNE
NPNVLEWLTPLKFIQISQEQMVSAGELFDPDIEVLKDLFCNEEGTYFPSPSVFTSPDIL
HSLRQIGLKNEASLKEKDVVQVAKKIEALQVGACPDQDVLLKKAKTLLLVLNKNHTLL
QSSEGKMTLKKIKWVPACKERPPNYPGSLVWKGDLNLCAPPDMCDVGHAILIGSSLP
LVESIHVNLEKALGIFTKPSLSAVLKHFKIVVDWYSSKTFSDDEYYQFHILLEIYGF
MHDHLNEGKDSFRALKFPWWVTGKKFCPLAQAVIKPIHDLDLQPYLHNPVKTMKFHQ
LFKVCGSIEELTSDHISMVQIKIYLKSDQDLSEQESKQNLHMLNIRWLYSNQIPAS
PNTVPVPIHHSKNPSKLIMKPIHECCYCDIKVDDLNDLLEDSEPIILVHEDIPMKTAE
WLKVPCLSTRLINPENMGFEQSGQREPLTVRIKNILEEYPSVSDIFKELLQONADDANA
TECSFLIDMRRNMDIRENLLDPGMAACHGPALWSFNNSQFSDSDFVNITRLGESLKR
EVDKVGKFGGLGFNSVYHITDIPIMREFMIMFDPNINHISKHIKDKSNPGIKINWSK
QQKRLRKFPNQFKPFIDVFGCQLPLTVEAPYSYNGTLFRLSFRTQQEAKVSEVSSTCY
NTADIYSLVDEFSLCGHLRIIFTQSVKSMYLYLKIEETNPSLAQDTVIIKKKSCSSK
ALNTPVLSVLKEAAKLMKTCSSSNKKLPSDEPKSSCILQITVEEFHHVFRIADLQSP
LFRGPDDPAALFEMAKSGQSKKPSDELSQKTVECTTWLLCTCMDTGEALKFSLSESG
RRLGLVPCGAVGVQLSEIQDQKWTVKPHIGEVFCYLPRIKTGLPVHINGCFAVTSNR
KEIWKTDTKGRWNTTFMRHVIVKAYLQVLSVLRDLATSGELMDYTYAVWPDPLVHD
DFSVICQGFYEDIAHGKGKELTKVFSKGSTWVSMKNVRFLDDSILKRRDVGSAAFKIF
LKYLKKTGSKNLCVELPSSVKLGFEAGCKQILLENTFSEKQFFSEVFFPNIQEIEA
ELRDPLMIFVLNEKVDEFSGVLRVTPCIPCSLEGHPLVLPRLIHPEGRVAKLFDIKD
GRFPYGSTQDYLNPILIKLVQLGMAKDDILWDDMLERAVSVAEINKSDHVAACLRSS
ILLSLIDEKLRKIRDPRAKDFAAKYQTIRFLPFLTKPAGFSLDWKGNFVKPETMFAATD
LYTAEHQDIVCLLQPIILNENSHSFRGCGSVSLAVKEFLGLLKKPTVDLVINQLKEVAK
SVDDGITLYQENITNACYKYLHEALMQNEITKMSIIDKLKPFSLVENAYVDSEKVS
FHLNFEAAPYLYQLPNKYKNNFRELFTVGVRSCTVEDFALVLESIDQERGKQITE
ENFQLCRRIISEGIWLSLIREKKQEFCEKNYKILLPDTNLMLLPAKSLCYNDPCWIKV

Figure 9A

KDTTVKYCHADIPREVAVKLGAVPKRHKALERYASNVCFTTLGTEFGQKEKLTSRIS
 ILNAYPSEKEMLKELLQNADDAKATEICFVFDPRQHPVDRI FDDKWAPLQGPALCVYN
 NQPFTEDDVRGIQNLGKGTKEGNPYKTGQYGIGFNSVYHITDCPSFISGNDILCIFDP
 HARYAPGATSISPGRMFRDLADFRTOFSDVLDLYLGTHFKLDNCTMFRFPLRNAEMA
 KVSEISSVPASDRMVQNLLDKLRSDGAELLMFLNHMEKISICEIDKSTGALNVLYSVK
 GKITDGDRLKRQFHASVIDSVTKKRQLKDIPVQQITYTMDTEDSEGNLTWILICNRS
 GFSSMEKVSXSIVSAHKNQDITLFPRGGAACITHNYKKPHRAFCFLPLSLETGLPFH
 VNGHFALDSARRNLWRDDNGVGRSDWNNSLMTALIPAYVELLIQLKKRYFPGSDPT
 LSVLQNTPIHVVKDTLKKFLSFFPVNRLDLQPDLYCLVKALYNCIHEDMKRLLPVVRA
 PNIDGSDLHSAVITWINMSTSNKTRPFFDNLLQDELQHLKNADYNITTRKTVAENVY
 RLKHLLEIGFNLVYNCDATANLYHCLIDADIPVSYPVPADIRSFLMTFSSPDTNCHI
 GKLPCLRLQQTNLKLFHSLKLLVDYCFKDAEENEIEVEGLPLLITLDSVLQTFDAKRPK
 FLTTYHELIPSRKDLFMENTLYLKYSNILLNCKVAKVFDISSFADLLSSVLPREYKTKS
 CTKWKDNFASESWLKNAWHFISESVSVKEDQEETKPTFDIVVDTLKD WALLPGTKFTV
 SANQLVVPEDVLLPLSLMHIAVFPNAQSDKV FHALMKAGCIQLALNKICSKDSAFVP
 LLSCHTANIESPTSILKALHYMVQTSTFRAEKLVENDFEALLMYFNCNLNHLMSQDDI
 KILKSLPCYKSISGRYVSIGKFGTCYVLTKSIPSAEVEKWTQSSSAFLEEKIHLKEL
 YEYIGCVPVDDLEVLYLHLLPKIENLSYDAKLEHLIYLKNRLSSAEELSEIKEQLFEK
 LESLLIHDANSRLKQAKHFYDRTVRVFEVMLPEKLFIPNDFFKLEQLIKPNHVT
 MTSWVEFLRNIGLYILSQQLLQFAKEISVRANTENWSKETLQNTVDILLHHIFQER
 MDLLSGNFKELSLIPFLC PERAPAEFIRFHPQYQEVNGTLPLIKFNGAQVNP KFKQC
 DVLQLLWTS CPILPEKATPLSIKEQEGSDLGPQEQLQVLNMLNVNLDPPLDKVINNC
 RNICNITTLDEEMVKTRAKVLRSIYEFLSAEKREFRQLRGVAFVMVEDGWKLLKPEE
 VVINLEYESDFKPYLYKLPLELGTFFHQLFKHLGTEDIISTKQYVEVLSRIFKNSEKQ
 LDPNEMRTVKRVVSGLFRSLQND SVKVRSDLENVRDLALYLP SQDGRLVKSSILVFDD
 APHYKSRIQGNIGVQMLVDLSQCYLGKDHGFHTKLIMLFPQKLRPRLSSILEEQ LDE
 ETPKVCQFGALCSLQGRQLLLSSEQFITGLIRIMKHENDNAFLANEKAI RLCALR
 EGLKVSCFEKLQTTLRVKGFNP IPHSRSETFAFLKRFGNVILLYIQHSDSKDINFL
 ALAMTLKSATDNLISDTSYLIAMLG CNDIYRIGEKLD SLGVKYDSSEPSKLELPMPGT
 PIPAEIHYTLLMDPMNVFPYGEYVGYLVDAEGGDIYGSYQPTYTYAIIVQEVEREDAD
 NSSFLGKIYQIDIGYSEYKIVSSLDLYKFSRPEESSQSRDSAPSTPTSPTEFLTPGLR
 SIPPLFSGRESHKTSKHKQSPKKLVNSLPEILKEVTSVVEQAWKLPESEK KIIIRRL
 YLKWHPDKNPENHDIANEVFKHLQNEINRLEKQAFLDQ NADRASRRTFSTSASRFQSD
 KYSFORFYTSWNQEATSHKSERQQQNK EKCPPSAGQTYSQRFVPPTFKSVGNPVEAR
 RWLRQARANFSAARNDLHKNA NEWVCFKCYLSTKLALIAADYAVRGKSDKDVKPTALA
 QKIEEYSQGLEGLTNDVHTLEAYGVDSLKTRYPDLLPFPQIPNDRFTSEVAMRVMECT
 ACIIIKLENFMQQKV"

BASE COUNT 4163 a 2256 c 2487 g 3887 tORIGIN

atgatttaca ggaagaccat gtactcagct gcagcttcta aatccagaac gatttgcacg
 tcttatcaag gaagtaatga atacattctg gcctggcaga gaattgattg ttcaatggta
 tccatttgat gaaaacagaa atcacccatc tgtttcatgg cttaagatgg tttggaaaaa
 tctttatata cttttttcag aggatttgac tttatttgat gagatgccac ttatccccag
 aactatacta gaggaaggtc agacatgtgt ggaactcatt agactcagga ttccatcggt
 agtcatttta gacgatgaat ctgaagcaca gcttccagaa ttttttagcag acattgtaca
 aaaacttgga ggggttggtc ttaaaaaatt agatgcatct atacaacatc cgcttattaa
 aaaatatatt cattcaccat taccaagtgc tgttttgagc ataattggaga agatgccatt
 gcagaaattg tgtaatcaaa taacttcgct acttccaaca cacaaagatg ccctgaggaa
 gttcttggt agtttaaccg atagcagtga gaaagagaaa agaattattc aagaattggc
 aatattcaag cgcattaacc attcttctga tcagggaatt tcctcttata caaaattgaa
 aggttgtaaa gtcttacacc atactgccaa actcccagca gatctgagac tttctatttc
 agtaatagac agtagtgatg aagctactat tcgtctggca aacatgttga aaatagaaca
 gttaaagacc actagctgct taaagcttgt tttaaaagat attgaaaatg cattttattc
 acatgaagag gtaacacagc ttatgttatg ggtccttgag aatctatctt ctcttaaaaa

Figure 9B

tgagaatcca	aatgtgcttg	agtgggttaac	accattaaaa	ttcatccaga	tatcacagga
acagatggta	tcagctgggtg	aactcttttga	ccctgatata	gaagtactaa	aggatctctt
ttgtaatgaa	gaaggaacct	atttcccacc	ctcagttttt	acctcaccag	atattctttca
ctccttaaga	cagattgggtt	taaaaaacga	agccagtctc	aaagaaaagg	atgttggtgca
agtggcaaaa	aaaattgaag	ccttacaggt	cggtagcttg	cctgatcaag	atgttcttct
gaagaaagcc	aaaacctct	tactggtttt	aaataagaat	cacacactgt	tgcaatcatc
tgaaggaaag	atgacattga	agaaaataaa	atgggttcca	gcttgcaagg	aaaggcctcc
aaattatcca	ggctcttttg	tctggaaagg	agatctctgt	aatctctgtg	caccaccaga
tatgtgtgat	gtaggccatg	caattctcat	tggctcctca	cttctctctg	ttgaaagtat
ccatgtaaac	ctggaaaaag	cattagggat	cttcacaaaa	cctagcctta	gtgctgtctt
aaaacacttt	aaaattggtg	ttgattggta	ttcttcaaaa	accttttagtg	atgaagacta
ctatcaattc	cagcatattt	tgcttgagat	ttacggattc	atgcatgac	atctaaatga
agggaaagat	tcttttagag	ccttaaaatt	tccatgggtt	tggactggca	aaaagttttg
tccacttgcc	caggctgtga	ttaaaccaat	ccatgatctt	gaccttcagc	cttatttgca
taatgtacct	aaaaccatgg	caaaattcca	ccaactattt	aaggctctgtg	gttcaataga
ggagttgaca	tcagatcata	tttccatggg	tattcagaag	atatatctca	aaagtgacca
agatctcagt	gaacaagaaa	gcaaacaaaa	tcttcatctt	atgttgaata	ttatcagatg
gctgtatagc	aatcagattc	cagcaagccc	caacacacca	gttctctatac	atcatagcaa
aaatccttct	aaacttatca	tgaagccaat	tcacgaatgc	tgttattgtg	acattaaagt
tgatgacctt	aatgacttac	ttgaagattc	tgtggaacca	atcatttttg	tgcattgagga
catacccatg	aaaactgcag	aatggctaaa	agttccatgc	cttagtaca	gactgataaa
tctgaaaac	atgggatttg	agcagtcagg	acaaagagag	ccacttactg	taagaattaa
aaatattctg	gaagaatacc	cttcagtgtc	agatattttt	aaagaactac	ttcaaaacgc
tgatgatgca	aatgcaacag	aatgcagttt	cttgattgat	atgagaagaa	atatggacat
aagagagaat	ctcctagacc	cagggatggc	agcttgtcat	ggacctgctt	tgtgggtcatt
caacaattct	caattctcag	attcagattt	tgtgaacata	actaggtttag	gagaatcttt
aaaaagggga	gaagttgaca	aagttggaaa	atttggctct	ggatttaatt	ctgtgtacca
tatcactgac	attcccatca	ttatgagtcg	ggaattcatg	ataatgttcg	atccaaacat
aaatcatatc	agtaaacaca	ttaaagacaa	atccaatcct	gggatcaaaa	ttaattggag
taaacaacag	aaaagactta	gaaaatttcc	taatcagttc	aaaccattta	tagatgtatt
tggctgtcag	ttacctttga	ctgtagaagc	accttacagc	tataatggaa	cccttttccg
actgtccttt	agaactcaac	aggaagcaaa	agtgagtga	gttagtagta	cgtgctacaa
tacagcagat	atttattctc	ttgtggatga	atttagtctc	tgtggacaca	ggcttatcat
tttcactcag	agtgtaaagt	caatgtattt	gaagtacttg	aaaattgagg	aaaccaaccc
cagtttagca	caagatacag	taataattaa	aaaaaaatcc	tgtcttcca	aagcattgaa
cacacctgtc	ttaagtgttt	taaaagaggc	tgttaagctc	atgaagactt	gcagcagcag
taataaaaaag	cttcccagtg	atgaacaaaa	gtcatcttgc	attcttcaga	tcacagtggg
agaatttcac	catgtgttca	gaaggattgc	tgattttacag	tcgccacttt	ttagagggtcc
agatgatgac	ccagctgtct	tctttgaaat	ggctaagctc	ggccaatcaa	aaaagccatc
agatgagttg	tcacagaaaa	cagtagagtg	taccacgtgg	cttctgtgta	cttgcatgga
cacaggagag	gctctgaagt	tttccctgag	tgagagtggg	agaagactag	gactgggtcc
atgtggggca	gtaggagttc	agctgtcaga	aatccaggac	cagaagtggg	cagtgaacc
acacattgga	gagggtgttt	gctattttacc	tttacgaata	aaaacaggct	tgccagttca
tatcaatggg	tgctttgctg	ttacatcaaa	taggaaagaa	atctggaaaa	cagatacaaa
aggacgatgg	aataccacgt	tcatgagaca	tgttattgtg	aaagcttact	tacagggtact
gagtggtctta	cgggacctgg	ccactagtgg	ggagctaatg	gattatactt	actatgcagt
atggcccgat	cctgatttag	ttcatgatga	tttttctgta	atttgccaag	gattttatga
agatatagct	catggaaaaag	ggaaagaact	gaccaaagtc	ttctctgatg	gatctacttg
ggtttccatg	aagaacgtaa	gatttctaga	tgactctata	cttaaaaagaa	gagatgttgg
ttcagcagcc	ttcaagatat	ttttgaaata	cctcaagaag	actgggtcca	aaaacctttg
tgctgttgaa	cttctctctt	cggtaaaaatt	aggatttgaa	gaagctggct	gcaaacagat
actacttgaa	aacacatttt	cagagaaaaca	gtttttttct	gaagtgtttt	ttccaaatat
tcaagaaatt	gaagcagaac	ttagagatcc	tttaatgatc	tttgttctaa	atgaaaaagt
tgatgagttc	tcgggagttc	ttcgtgttac	tccatgtatt	ccttgttcct	tggaggggca

Figure 9C

tcctttggtt	ttgccatcaa	gattgatcca	ccccgaagga	cgagttgcaa	agttatttga
tattaaaagat	gggagattcc	cttatggttc	tactcaggat	tatctcaatc	ctattatttt
gattaaaacta	gttcagttag	gtatggcaaa	agatgatatt	ttatgggatg	atatgctaga
acgtgcagtg	tcagtagctg	aaattaataa	aagtgatcat	gttgctgcat	gcctaagaag
tagtatctta	ttgagtctta	tcgatgagaa	actaaaaata	agggatccta	gagcaaagga
ttttgctgca	aaatatcaaa	caatccgctt	ccttccattt	ctgacaaaac	cagcagggtt
ttctttggac	tggaaaaggca	acagttttta	gcctgaaacc	atgtttgcag	caactgacct
ttatacagct	gaacatcaag	atatagtttg	tcttttgcaa	ccaattctaa	atgaaaattc
ccattctttt	agaggtttg	gttcagtgtc	attggctgtt	aaagagtttt	tgggattact
caagaagcca	acagttgatc	tggttataaa	ccaattgaaa	gaagtagcaa	aatcagttga
tgatggaaat	acactgtacc	aggagaatat	caccaatgct	tgctacaaat	accttcatga
agccttgatg	caaaatgaaa	tactaagat	gtcaattatt	gataagttaa	aaccctttag
cttcattcta	gttgagaatg	catatgttga	ctcagaaaag	gtttcttttc	atttaaattt
tgaggcggca	ccataccttt	atcagttgcc	taataagtat	aaaaataatt	tccgcgaact
ttttgaaacc	gtgggtgtga	ggcagtcag	cactgttgaa	gattttgctc	ttgttttggg
atctattgat	caagaaagag	gaacaaagca	aataacagaa	gagaattttc	agctttgccg
acgaataatc	agtgaaggaa	tatggagtct	cattagagaa	aagaaacaag	aattttgtga
gaaaaattat	ggcaagatat	tattgccaga	tactaatctt	atgcttctcc	ctgctaaatc
gttatgctac	aatgattgcc	cttggataaa	agtaaaggat	accactgtaa	aatattgtca
tgctgacata	cccagggaag	tagcagtaaa	actaggagca	gtcccaaagc	gacacaaagc
cttagaaaaga	tatgcatcca	atgtctgttt	tacaacactt	ggcacagaat	ttgggcagaa
agaaaaattg	accagcagaa	ttaagagcat	ccttaatgca	tatccttctg	aaaaggaaat
gttgaaaagag	cttcttcaaa	atgctgatga	tgcaaaggcg	acagaaatct	gttttgtgtt
tgatcctaga	cagcatccag	ttgatagaat	atttgatgat	aagtgggccc	cattgcaagg
gccagcactt	tgtgtgtaca	acaaccagcc	atttacagaa	gatgatgtta	gaggaattca
gaatcttgga	aaaggcacga	aagagggaaa	tccttataaa	actggacagt	atggaatagg
attcaattct	gtgtatcata	tcacagactg	cccatctttt	atttctggca	atgacatcct
gtgtattttt	gatcctcatg	ccagatatgc	accagggggc	acatccatta	gtcccggacg
catgtttaga	gatttggatg	cagatttttag	gacacagttc	tcagatgttc	tggatcttta
tctgggaacc	catttttaaac	tggataattg	cacaatgttc	agatttcctc	ttcgtaatgc
agaaatggca	aaagtttcgg	aaatttcgtc	tgttccagca	tcagacagaa	tgggtccagaa
tcttttggac	aaactgcgct	cagatggggc	agaacttcta	atgtttctta	atcacatgga
aaaaatttct	atttgtgaaa	tagataagag	tactggagct	ctaaatgtgc	tgtattcagt
aaagggcaaa	atcacagatg	gagacagatt	gaaaaggaaa	caatttcatg	catctgtaat
tgatagtgtt	actaaaaaga	ggcagctcaa	agacatacca	gttcaacaaa	taacctatac
tatggatact	gaggactctg	aaggaaatct	tactacgtgg	ctaatttgta	atagatcagg
cttttcaagt	atggagaaaag	tatctaaaag	tgtcatatca	gctcacaaga	accaagatat
tactcttttc	ccacgtggtg	gagtagctgc	ctgcattact	cacaactata	aaaaacccca
tagggccttc	tgttttttgc	ctctttcttt	ggagactggg	ctgccatttc	atgtgaatgg
ccactttgca	ctggattcag	ccagaaggaa	cctgtggcgt	gatgataatg	gagttggtgt
tcgaagtgac	tgggaataaca	gtttaatgac	agcattaata	gctcctgcat	atgttgaatt
gctaatacag	ttaaaaaaac	ggtatttccc	tggttctgat	ccaacattat	cagtgttaca
gaacacccct	attcatgttg	taaaggacac	tttaaagaag	tttttatcgt	ttttcccagt
taaccgtctt	gatctacagc	cagatttata	ttgtctagtg	aaagcacttt	acaattgcat
tcacgaagac	atgaaacgtc	ttttacctgt	tgtgcgggct	ccaaatattg	atggctctga
cttgcaactct	gcagttataa	ttacttggat	caatatgtct	acttctaata	aaactagacc
attttttgac	aatttactac	aggatgaatt	acaacacctt	aaaaatgcag	attataatat
caccacacgc	aaaacagtag	cagagaatgt	ctataggctg	aaacatctcc	ttttagaaat
tggtttcaac	ttggtttata	actgtgatga	aactgcta	ctttaccact	gtcttataga
tgacagatatt	cctgttagtt	atgtgacccc	tgctgatata	agatcttttt	taatgacatt
ttcctctcct	gacactaatt	gccatattgg	gaagctgcct	tgtcgtctgc	agcagactaa
tctaaaactt	tttcatagtt	taaaactttt	agttgattat	tgttttaaag	atgcagaaga
aaatgagatt	gaagttgagg	gattgcccct	tctcatcaca	ctggacagtg	ttttgcaaac
ttttgatgca	aaacgaccca	agtttcta	aacatatcat	gaattgattc	catcccgc

Figure 9D

cactgagttc	ctcacccttg	gcttgagaag	cattcctcct	cttttctctg	gtagagagag
ccacaagact	tcttccaaac	atcagtcccc	caaaaagctt	aagggttaatt	ctttaccaga
aatcttaaaa	gaagtgacat	ctgtggtgga	gcaagcatgg	aagcttccag	aatcggaacg
aaaaaagatt	attaggcggt	tgtatttgaa	atggcatcct	gacaaaaatc	cagagaacca
tgacattgcc	aatgaagttt	ttaaacattt	gcagaatgaa	atcaacagat	tagaaaaaca
ggctttttcta	gatcaaaaatg	cagacagggc	ctccagacga	acattttcaa	cctcagcatc
ccgattttcag	tcagacaaaat	actcattttca	gagatttctat	acttcatgga	atcaagaagc
aacgagccat	aaatctgaaa	gacagcaaca	gaacaaagaa	aaatgcccc	cttcagccgg
acagacttac	tctcaaagggt	tctttgttcc	tcccactttc	aagtcgggtg	gcaatccagt
ggaagcacgc	agatggctaa	gacaagccag	agcaaaacttc	tcagctgcca	ggaatgacct
tcataaaaaat	gccaatgagt	gggtgtgctt	taaatgttac	ctttctacca	agtttagcttt
gattgcagct	gactatgctg	tgaggggaaa	gtctgataaa	gatgtaaaac	caactgcact
tgctcagaaa	atagaggaat	atagtcagca	acttgaagga	ctgacaaatg	atgttcacac
attggaagct	tatggtgtag	acagttttaa	aacaagatac	cctgatttgc	ttccctttcc
tcagatccca	aatgacaggt	tcacttctga	ggttgctatg	aggggtgatgg	aatgtactgc
ctgtatcata	ataaaaacttg	aaaattttat	gcaacaaaaa	gtgtgaagat	atttaacgaa
aaaaaaggta	gatcttgaat	gtgttgtagc	acgaataaat	tgctgtactt	cattaagctt
cattgccaat	tagctaggaa	ttgttaagca	cattgcagat	tgttcttggg	gaattctgga
gttggttatga	acatgaatac	caacggaaaa	ccttaactga	atctaaaaga	aaactatttt
gaagatggtg	gtgagctgca	aaatagctgg	atggatttga	atgattggga	tgatacatca
ttgaactgca	ctttatataa	ccaaagctta	gcagtttggt	agataagagt	ctatgtatgt
ctctgggttag	gatgaagtta	attttatggt	tttaacatgg	tatttttgaa	ggagctaattg
aaacactgga	catataattg	gtttaaacat	aagggggaatt	aagtccttgt	agtctgtcat
ttttttaagt	ggatcctctt	ggatgcgtta	ttttctcatc	agctggctct	gatcatgaat
ttgttgtaat	tttatgttgt	actcagtgca	tttaagaaat	ggtagagtat	tttaatccta
ttacttgact	aagagtgtga	aggtagtact	ttttagagtg	cactgagtgc	actttacatc
tttattttaa	ttttttttta	acatcttatg	tttacaggct	tctgttttga	tgaagatagc
aacggaaaaac	tcaaaatggt	ggcagttcct	attaccagtt	gttagtattg	tttctggaaa
ctgcttgcca	agacaacatt	tattaactgt	tagaacactt	gctttatggt	tgtgtgtaca
tattttccac	aaatgttata	atttatatag	tgtggttgaa	caggatgcaa	tcttttggtg
tctaaagggtg	ctgcagttaa	aaaaaaaaaca	accttttctt	tcaatatggc	atgtagtgga
gttttttttaa	ctttaaaaaac	atcaaaaaatt	gttaaaatca	ttgtgttatc	tagtagttta
taattatcgg	cttatatttc	cccatgaatg	atcagaactg	acatttaatt	catgtttgtc
tcgccatgct	tctttacttt	aacatatttc	ttttgcagaa	tgtaaaagggt	aatgataatt
agttttatata	agtgtactgg	ctgtaaaatga	tgctaaaatat	acttttatgca	attaagggtc
tacagaacat	gttgaaaactt	tttttacttt	tattgggaat	aagggaatggt	tgcacctcca
cattttattg	ctt				

Figure 9F